View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by INRIA a CCSD electronic archive server

Discrete minimum distortion correspondence problems for non-rigid shape matching

Chaohui Wang, Michael M. Bronstein, Nikolaos Paragios

To cite this version:

Chaohui Wang, Michael M. Bronstein, Nikolaos Paragios. Discrete minimum distortion correspondence problems for non-rigid shape matching. [Research Report] RR-7333, INRIA. 2010. inria-00498591

HAL Id: inria-00498591

https://hal.inria.fr/inria-00498591

Submitted on 7 Jul 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

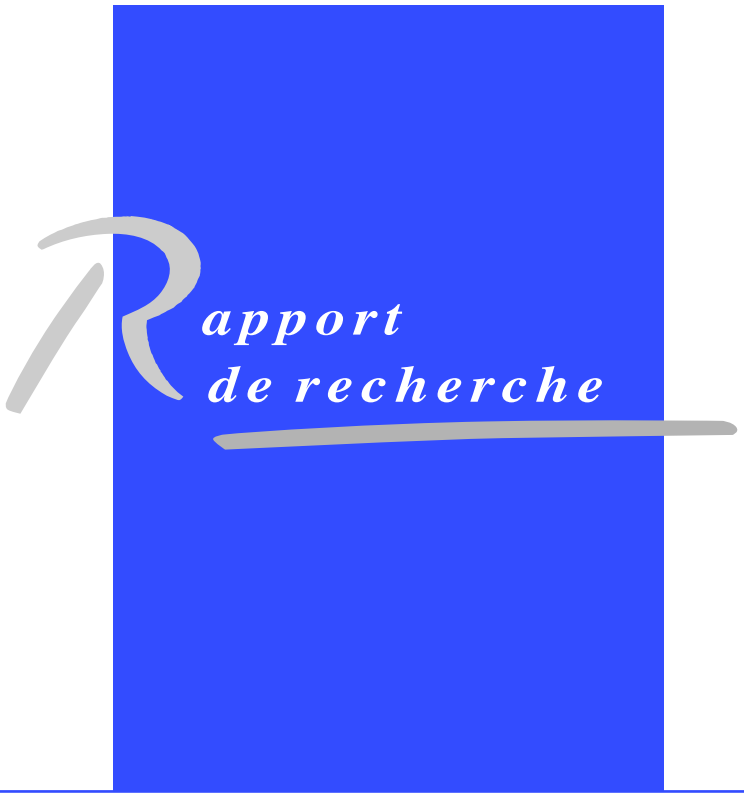
INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Discrete minimum distortion correspondence problems for non-rigid shape matching

Chaohui Wang — Michael M. Bronstein — Nikos Paragios

N° 7333

July 2010



ISSN 0249-6399 ISRN INRIA/RR--7333--FR+ENG

INSTITUT NATIONAL
DE RECHERCHE
EN INFORMATIQUE
ET EN AUTOMATIQUE

INRIA

centre de recherche
SACLAY - ÎLE-DE-FRANCE

Discrete minimum distortion correspondence problems for non-rigid shape matching

Chaohui Wang*† , Michael M. Bronstein‡, Nikos Paragios*†

Thème : reconaissance de formes
Équipe-Projet Gallen

Rapport de recherche n° 7333 — July 2010 — 19 pages

Abstract: Similarity and correspondence are two fundamental archetype problems in shape analysis, encountered in numerous application in computer vision and pattern recognition. Many methods for shape similarity and correspondence boil down to the minimum-distortion correspondence problem, in which two shapes are endowed with certain structure, and one attempts to find the matching with smallest structure distortion between them. Defining structures invariant to some class of shape transformations results in an invariant minimum-distortion correspondence or similarity. In this paper, we model shapes using local and global structures and formulate the invariant correspondence problem as binary graph labeling. We perform challenging non-rigid shape matching experiments, and show how different choice of structure results in invariance under various classes of deformations.

Key-words: shape recognition, shape similarity, Gromov-Hausdorff metric

* Laboratoire de Mathématiques Appliquées aux Systèmes, École Centrale de Paris, France
† Equipe GALEN, INRIA Saclay - Île de France, Orsay, France
‡ Department of Computer Science, Technion – Israel Institute of Technology

Centre de recherche INRIA Saclay – Île-de-France
Parc Orsay Université
4, rue Jacques Monod, 91893 ORSAY Cedex
Téléphone : +33 1 72 92 59 00

Problèmes discret de correspondance de distorsion minimale pour le recalage de formes déformables

Résumé : La similarité et la correspondance sont deux problèmes archetypaux fondamentaux dans le cardre de l'analyse de formes, rencontrés dans de nombreuses applications en vision par ordinateur et reconnaissance de formes. De nombreuses méthodes approchant la similarité et la correspondance se resument à un problème de correspondance de distorsion minimale, dans lequel deux formes sont dotées d'une certaine structure, la correspondance avec distorsion structurelle minimale est alors recherchée. La définition de structures invariantes dans certaines classes de tranformations de formes aboutit à une correspondance ou à une similarité de distorsion minimale invariante. Dans ce rapport, nous modélisons des formes en utilisant des structures locales et globales, et nous formulons le problème de correspondance invariante sous forme de graphe binaire. Nous effectuons une série d'expériences ambitieuses sur la correspondance de formes déformables, et nous montrons comment les différents choix de structure aboutissent à l'invariance sous différentes classes de déformations.

Mots-clés : reconaissance de formes, similarité des formes, distance de Gromov-Hausdorff

1 Introduction

Recent works in computer vision and shape analysis [2, 31, 8, 9] have shown that different approaches to shape similarity and correspondence can be considered as instances of the *minimum distortion correspondence problem*, in which two shapes are endowed with certain structure, and one attempts to find the best (least distorting) matching between these structures. Examples of such structures include multiscale heat kernel signatures [43, 34, 6, 14], local photometric properties [48, 45], topological graphs [4], conformal factor [1], shape extremities [50, 51], eigenfunctions of the Laplace-Beltrami operator [40, 28, 22, 16], triplets of points [26, 49], and geodesic [31, 8, 17], diffusion [11, 30] and commute time [40, 36, 12, 13] distances. By defining a structure invariant under certain class of transformations (e.g. non-rigid deformations), one obtains correspondence invariant under that class (in the above example, deformation invariant matching). The Gromov-Hausdorff distance [20] is an important particular case of the minimum distortion correspondence problem, in which the matched structures are metric spaces, invariant to isometries of the metric structures.

Some settings of the minimum distortion correspondence problem can be reformulated as labeling problems [47, 42], such that the objective function can be optimized efficiently using the recently developed discrete optimization algorithms. For example, the dual-decomposition strategy [3], introduced by [24] to perform pairwise Markov random field (MRF) inferences, provides a powerful technique to solve such labeling problems. Based on such a strategy, Torresani *et al.* proposed a pairwise graph matching algorithm [47] to compute correspondence between images using a criterion combining local features and Euclidean distances between nearby features. Such an approach showed better performance than feature-only based methods in deformable 2D object tracking, the increased performance attributed to the use of inter-feature distances as a geometric consistency constraint. However, Euclidean distances are not deformation invariant and can be applied only locally, thus limiting the usefulness of such a constraint.

1.1 Main contribution

In this paper, we study the minimum distortion correspondence problem in the context of non-rigid shape analysis. We formulate *invariant correspondence* as a minimizer of a distortion criterion based on structures invariant to some classes of transformations. In particular, we use local and global structures invariant to important classes of transformations such as non-rigid deformations, changes in topology, and scaling. By such an axiomatic construction of invariant structures, we obtain invariant correspondence. In particular, we show scale invariant shape matching using only singleton and pairwise interactions without higher-order terms. Compared to Torresani *et al.* [47], our use of global structures in non-rigid shapes provides a better regularization to the problem and is better motivated geometrically. Yet, it also increases the computational complexity of the optimization. To address this problem, we use hierarchical matching, in which candidate correspondences are restricted to neighborhoods of matching points from coarser levels.

While the described axiomatic approach is suitable for modeling geometric shape transformations such as bendings, it is not applicable to intra-class shape variations (e.g. different appearances of a human shape). To cope with this case, we show a probabilistic extension of our framework, in which local and global structures are replaced with respective multidimensional distributions, accounting for shape variability.

1.2 Related work

Feature-based shape matching methods for non-rigid shapes were used in numerous recent works [48, 43, 22]. Tree-based [51], RANSAC-type [44], and branch-and-bound techniques [37, 38, 39] were used to find the matches between the feature points. Elad and Kimmel [17] used multidimensional scaling (MDS) [5] to represent shapes in a low-dimensional Euclidean space and compare them as rigid objects. The use of an intermediate embedding space was eliminated in [31] using the Gromov-Hausdorff formalism [20]. Bronstein *et al.* [8] proposed an MDS-like algorithm referred to as generalized MDS (GMDS) for the computation of the Gromov-Hausdorff distance and deformation invariant correspondence between shapes [9]. This framework was extended in [11, 30] using diffusion geometry instead of the geodesic one. In [29], Mémoli extended [31, 8] by modeling shapes as metric-measure spaces. He introduced the Gromov-Wasserstein distance based on measure coupling between two metric-measure spaces, and formulated is as a quadratic assignment problem (QAP). Thorstensen and Keriven [45] extended the GMDS framework to textured shapes introducing *photometric stress* as a local matching term in addition to geodesic distance distortion. Dubrovina and Kimmel [16] generalized this approach for the matching of textureless shapes using Laplace-Beltrami eigenfunctions as local geometric descriptors. Mateus *et al.* [28] showed a non-rigid shape correspondence approach with inexact graph matching based on spectral embedding. In the image domain, Torresani *et al.* [47] used graph labeling problem to match 2D images.

2 Problem formulation

Our shape model is an extension of the *metric model* used in [31, 8, 17]. We assume that the shapes are endowed with *local* and *global* structure, and try to find such a correspondence between the shapes that best preserves these structures. The structures are defined having in mind certain invariance properties required in the particular problem, as discussed in Section 3. Given a shape X, modeled as a connected surface (possibly with boundary) embedded into $\mathbb{R}^3$ (or $\mathbb{R}^2$ in case of planar shapes), its local structure is modeled by a vector field $\mathbf{f}_X : X \to \mathbb{R}^m$ referred to as a *local descriptor*. For example, $\mathbf{f}_X$ can be texture [47], local photometric [48] or geometric [33, 46, 1, 43, 34, 14] descriptor. The global structure of the shape is modeled as a *metric* $d_X : X \times X \to \mathbb{R}$, defined as a positive-definite subadditive function between pairs of points on X.

Given two shapes X and Y with the local descriptors $\mathbf{f}_X$ and $\mathbf{f}_Y$ and metrics d_X and d_Y, respectively, we define a *bijective correspondence* between X and Y as $\mathcal{C} \subset X \times Y$ satisfying $\forall x \in X\ \exists! y \in Y$ such that $(x, y) \in \mathcal{C}$ and $\forall y \in Y\ \exists! x \in X$ such that $(x, y) \in \mathcal{C}$. A good correspondence should match similar descriptors between corresponding points and similar metrics between corresponding pairs of points. This can be quantified using *first-* and *second-order distortion* terms, $\mathrm{dis}(\mathcal{C}) = \|\mathbf{f}(\mathcal{C})\|$ and $\mathrm{dis}(\mathcal{C} \times \mathcal{C}) = \|\mathbf{d}(\mathcal{C} \times \mathcal{C})\|$, measuring the quality of correspondence of local and global structures, respectively. (here, $\mathbf{f}(\mathcal{C})$ is a $|\mathcal{C}| \times 1$ vector with elements $\|\mathbf{f}_X(x) - \mathbf{f}_Y(y)\|$ for all $(x, y) \in \mathcal{C}$; $\mathbf{d}(\mathcal{C} \times \mathcal{C})$ is a $|\mathcal{C}|^2 \times 1$ vector with elements $\|d_X(x, x') - d_Y(y, y')\|$

for all $(x,y),(x',y') \in \mathcal{C}$; and $\|\cdot\|$ is some norm). In particular,

$$\begin{aligned}\mathrm{dis}_2(\mathcal{C}) &= \sum_{(x,y)\in\mathcal{C}} \|\mathbf{f}_X(x) - \mathbf{f}_Y(y)\|^2;\\ \mathrm{dis}_2(\mathcal{C}\times\mathcal{C}) &= \sum_{(x,y),(x',y')\in\mathcal{C}} (d_X(x,x') - d_Y(y,y'))^2.\end{aligned}$$

The optimal correspondence is found by minimizing a combination of first- and second-order distortion terms,

$$\min_{\mathcal{C}} \mathrm{dis}(\mathcal{C}) + \beta\mathrm{dis}(\mathcal{C}\times\mathcal{C}), \quad \beta \geq 0. \tag{1}$$

The minimizer of problem (1) is the *minimum distortion correspondence* between X and Y. The minimum of problem (1) can be interpreted as the similarity of X and Y. [1] A particular theoretically important case is a minimum-distortion correspondence with an L_∞ second-order distortion term, referred to as the *Gromov-Hausdorff distance* [20]

$$d_{\mathrm{GH}}(\mathcal{C}\times\mathcal{C}) = \frac{1}{2}\min_{\mathcal{C}} \max_{(x,y),(x',y')\in\mathcal{C}} |d_X(x,x') - d_Y(y,y')|.$$

3 Invariance

The choice of the local and global structures ($\mathbf{f}_X, \mathbf{f}_Y$ and d_X, d_Y) defines the invariance properties of the correspondence. Assume that the shape $Y = \tau(X)$ is obtained from X by means of some transformation τ from a class $\mathcal{T}$. If $\mathbf{f}_X \circ \tau = \mathbf{f}_Y$ and $d_X \circ (\tau\times\tau) = d_Y$ for all $\tau \in \mathcal{T}$, our structures are *invariant* under the class of transformations $\mathcal{T}$. As a result, correspondence obtained by the solution of problem (1) is also invariant under $\mathcal{T}$. Important invariance classes can be addressed by appropriate definition of the descriptors and the metric. In particular, we are interested in *inelastic deformations* (bendings), changing the embedding of the shape without changing its intrinsic structure; *topological transformations*, resulting in local changes in the connectivity of the shape, appearing as holes or "gluing" two points on the surface; and *scaling*.

3.1 Choice of the metric

3.2 Geodesic metric

One of the most straightforward definitions of a metric on a surface is the *geodesic metric*, measuring the length of a shortest path between points x and x',

$$d_X(x,x') = \min_{\gamma\in\Gamma(x,x')} \ell(\gamma),$$

where $\Gamma(x,x')$ denotes the set of all admissible paths between x and x', γ is some admissible path, and $\ell(\gamma)$ is its length. The geodesic metric is intrinsic, dependent only

[1] The minimizer of problem (1) is not necessarily unique, i.e., there may be two different correspondences $\mathcal{C} \neq \mathcal{C}'$ with the same distortion. Such situations are typical when the shapes have *intrinsic symmetries*. Intrinsic symmetry is manifested by the existence of a self-isometry of X with respect to the metric d_X, i.e., an automorphism $g : X \to X$ satisfying $d_X = d_X \circ (g\times g)$ [37, 35, 38, 39]. If g and h are intrinsic symmetries of X and Y, respectively, the distortion of the correspondences $\mathcal{C}$ and $\mathcal{C}\circ(g\times h)$ is identical and the minimum-distortion correspondence problem has multiple minimizers [9, 37].

on local distance structure of the shape, and is thus invariant to inelastic deformations [31, 8, 17]. A notable drawback of the geodesic distance is its sensitivity to topological transformations. Connectivity changes alter the admissible paths Γ (e.g., gluing the fingers of the hand creates new paths that have not existed before), and, since the geodesic distance takes the minimum over all path lengths, sometimes the change in the geodesic metric can be very significant.

3.3 Diffusion metric

A more robust definition of an intrinsic metric based on heat diffusion properties has been recently popularized by Lafon *et al.* [15]. Heat diffusion on manifolds is governed by the *heat equation* $\left(\Delta_X + \frac{\partial}{\partial t}\right) u = 0$, where u is the heat distribution and Δ_X is the positive semi-definite *Laplace-Beltrami operator* (LBO), which can be roughly thought of as a generalization of the Laplacian to non-Euclidean domains. The *heat kernel* $h_{X,t}(x,z)$ is the solution of the heat equation with a point heat source at point x at time $t = 0$. For compact manifolds, the Laplace-Beltrami operator has discrete eigendecomposition of the form $\Delta_X \phi_i = \lambda_i \phi_i$, where $\lambda_0 = 0, \lambda_1, ... \geq 0$ are eigenvalues and $\phi_0, \phi_1, ...$ are eigenfunctions. Using the eigenbasis of Δ_X, the heat kernel can be presented as

$$h_{X,t}(x,z) \quad = \quad \sum_{i=0}^{\infty} e^{-\lambda_i t} \phi_i(x) \phi_i(z). \tag{2}$$

A family of metrics

$$d^2_{X,t}(x,y) \quad = \quad \|h_{X,t}(x,\cdot) - h_{X,t}(y,\cdot)\|^2_{L_2(X)} = \sum_{i=1}^{\infty} e^{-2\lambda_i t} (\phi_i(x) - \phi_i(y))^2, \tag{3}$$

parameterized by the time scale t, is referred to as *diffusion metrics*. Diffusion metric is inversely related to the connectivity of points x and y by paths of length t. Unlike the geodesic distance which measures the length of the shortest path, the diffusion metric has an averaging effect over all paths connecting two points. As a result, diffusion metric is less sensitive to topology and connectivity changes [11]. With an appropriate selection of the time scale t, the effect of topological noise can be reduced [12, 13].

3.4 Commute-time metric

At the same time, the need to select the scale parameter is a disadvantage, as it depends on the shape scale. Moreover, the diffusion metric is not scale invariant, since scale change affects the eigenvalues λ_i and eigenfunctions ϕ_i. A different metric,

$$\delta^2_X(x,y) \quad = \quad \sum_{i=1}^{\infty} \frac{1}{\lambda_i} (\phi_i(x) - \phi_i(y))^2, \tag{4}$$

called the *commute time* (or *resistance* [19]) *distance*, is similar in its spirit to the diffusion metric, while being scale-invariant. The commute time metric measures the connectivity of points by paths of any length and is related to the expected time it takes a random walk initiating at point x go through point y and return to x. It is related to the diffusion distance by the following formula [36],

$$\delta^2_X(x,y) \quad = \quad 2 \int_0^{\infty} d^2_{X,t}(x,y) dt.$$

Structure	**Bending**	**Topology**	**Scale**
Photometric [48]	Yes*	Yes**	Yes
Local histogram [37]	Yes	No	No
Gaussian curvature	Yes	No	No
Conformal factor [1]	Yes	No	No
LBO eigenfunctions [22, 16]	Yes	No	Yes
HKS [43]	Yes	Approx	No
SI-HKS [14]	Yes	Approx	Yes
Geodesic metric [31, 8]	Yes	No	No
Diffusion metric [11]	Yes	Approx	No
Commute-time metric [36, 12, 13]	Yes	Approx	Yes

Table 1: Invariance properties of local (top rows) and global (bottom rows) structures. * Assuming uniform illumination. ** Only point-wise, assuming no averaging is done on texture.

Due to integration over t, δ_X does not involve scale selection and is invariant to scaling transformations [12, 13].

3.5 Choice of the descriptor

Similarly to our motivation in the selection of the metric, the choice of the local descriptor is also dictated by the desired invariance properties. Due to their locality, many types of descriptors are usually less susceptible to changes as a result of non-rigid deformations. However, some descriptors have explicit invariance properties by construction.

Several works (e.g. [33, 46]) successfully used rigid shape patches to compare non-rigid shapes. Raviv *et al.* [37] used local histograms of geodesic distances as local shape descriptors. Such descriptors are deformation invariant, however, suffer from all the problems of geodesic metrics (sensitivity to topological noise and scale). Hu and Hua [22] and Dubrovina and Kimmel [16] used the pointwise values of the Laplace-Beltrami eigenfunctions as local descriptors. Such descriptors are deformation-invariant, but sensitive to changes in topology.

3.6 Heat kernel signature

Sun *et al.* [43] introduced intrinsic descriptors based on multi-scale heat kernels, referred to as *heat kernel signatures* (HKS). The HKS is constructed at every point of the shape by considering the values of the heat kernel diagonal at multiple time scales, $\mathbf{f}_X(x) = (h_{X,t_1}(x,x), \ldots, h_{X,t_n}(x,x))$, where $t_1, \ldots, t_n$ are some time scale. The HKS is invariant to inelastic deformations and was also shown to be insensitive to topological transformations [34].

3.7 Scale-invariant heat kernel signature

The disadvantage of HKS is the lack of scale invariance. In a follow-up work, Bronstein and Kokkinos [14] introduced a scale-invariant modification of HKS, referred to as SI-HKS. The main idea is to sample the time scales logarithmically ($t = \alpha^\tau$) such that shape scaling corresponds to a scale-space shift. Such a shift is then undone by taking

the magnitude of the Fourier transform w.r.t. τ. The SI-HKS enjoys the invariance properties of HKS, while in addition also being scale-invariant.

3.8 Photometric descriptors

If in addition the shape is endowed with some photometric properties or texture, they can be used as a local descriptor. The advantage of photometric descriptor is its insensitivity to any geometric transformations. Recent works of Thorstensten and Keriven [45] used a *photometric stress* in an extension of the GMDS framework [8]. Zaharescu *et al.* [48] constructed SIFT-like [27] texture descriptors on shapes. Such descriptors were shown to be very robust in shape matching with significant topological noise.

4 Correspondence as a graph labeling problem

Our minimum-distortion correspondence problem can be formulated as a *binary labeling* problem with uniqueness constraints [47] in a graph with vertices defined as pairs of points and edges defined as quadruplets. More formally, let $\mathcal{V} = \{(x,y) : x \in X, y \in Y\} = X \times Y$ be the set of pairs of points from X and Y, and let $\mathcal{E} = \{((x,y),(x',y')) \in \mathcal{V} \times \mathcal{V}$ and $(x,y) \neq (x',y')\}$. Let $\mathcal{L} = \{0,1\}$ further denote the set of binary labels. We can represent a correspondence $\mathcal{C} \subseteq \mathcal{V}$ as binary labeling $u \in \mathcal{L}^{\mathcal{V}}$ of the graph $(\mathcal{V},\mathcal{E})$, as follows: $u_{x,y} = 1$ iff $(x,y) \in \mathcal{C}$ and 0 otherwise. When using L_2 distortions, the correspondence problem (1) can be reformulated as

$$\begin{aligned} \min_{u \in \mathcal{L}^{\mathcal{V}}} \quad & \sum_{(x,y)\in\mathcal{V}} u_{x,y}(\|\mathbf{f}_X(x) - \mathbf{f}_Y(y)\| - \gamma) + \\ & \beta \sum_{((x,y),(x',y'))\in\mathcal{E}} u_{x,y}u_{x',y'}|d_X(x,x') - d_Y(y,y')|^2 \\ & \text{s.t.} \sum_y u_{x,y} \leq 1 \ \ \forall x \in X; \quad \sum_x u_{x,y} \leq 1 \ \ \forall y \in Y. \end{aligned} \tag{5}$$

where $\gamma > 0$ is an occlusion term [47] to penalize unmatched points. We can choose a sufficiently large γ to ensure the bijective correspondence and the equivalence of the two problems 1 and 5.

In general, optimization of this energy is NP-hard [18]. Here, we adopt the graph matching algorithm [47] based on dual decomposition to perform the optimization of (5). The key idea of this approach is, instead of minimizing directly the energy (5) of the original problem, to maximize a lower bound on it by solving the dual to the linear programming (LP) relaxation of (5). This approaches demonstrate good global convergence behavior [24]. We first decompose the original problem, which is too complex to solve directly, into a series of sub-problems, each of which is smaller and solvable. After getting the solution of the sub-problems, we combine them using a projected-subgradient scheme to get the solution of the original problem. In the numerical experiments, following [47], we decompose problem (5) into a *linear subproblem*, a *maxflow subproblem* and a set of *local subproblems*.

4.1 Hierarchical matching

Assuming for simplicity $|X| = |Y| = N$, the number of vertices in the graph is $|\mathcal{V}| = N^2$ and the number of edges, assuming full connectivity, is $\mathcal{O}(N^4)$. The complexity

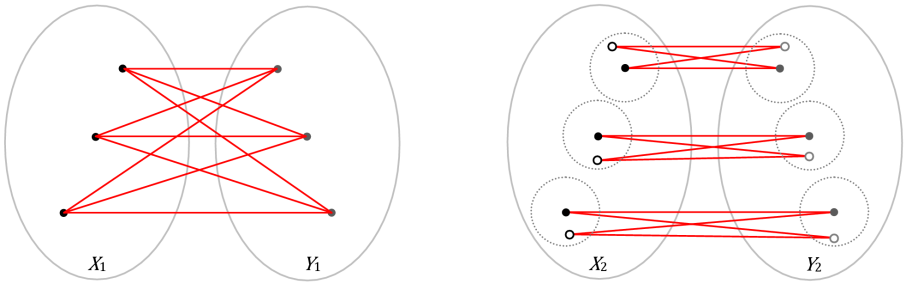


Figure 1: Toy example of hierarchical matching. Left: at the first level, $N_1 = 3$, $\mathcal{V}_1 = X_1 \times Y_1$ (all possible pairs of points). The graph size is $|\mathcal{V}_1| = 9$, $|\mathcal{E}_1| = 36$, counting only $(N_1^2 - N_1)/2$ distances. Right: at the second level, $N_2 = 6$ (3 points are added; $q = 2$). The graph size is $|\mathcal{V}_2| = 12$, $|\mathcal{E}_2| = 66$ (compare to full graph with $|\mathcal{V}_2| = 36$, $|\mathcal{E}_2| = 630$).

of problem (5) is $\mathcal{O}(|\mathcal{V}|^2|\mathcal{E}|)$ multiplied by the number of iterations, i.e., $\mathcal{O}(N^8)$. This complexity can be reduced by adopting a hierarchical matching strategy: after finding a coarse correspondence between a small number of points, correspondence between nearby points only is looked for. This allows to significantly reduce the graph size.

Let $x_1, x_2, \ldots$ denote a progressive sampling of the shape X, such that $X_n = \{x_1, \ldots, x_n\}$ constitutes an r_n-covering of X (i.e., $d_X(X, X_n) \leq r_n$, where d_X is some metric on X). Such a sequence of points can be found using e.g. farthest point sampling (FPS) strategy [21], in which x_1 is selected arbitrarily and the next point is selected as $x_{k+1} = \arg\max_{x \in X} \min_{i=1,\ldots,k} d_X(x, x_i)$. Same way, $Y_n = \{y_1, \ldots, y_n\}$ will denote an r'_n-covering of Y.

At the first stage of hierarchical matching, correspondence is found between X_{N_1} and Y_{N_1}, where N_1 is some small number (in our experiments, it varied between 4 and 10), solving the labeling problem (5) on the full graph ($\mathcal{V}_1 = X_{N_1} \times Y_{N_1}$, $\mathcal{E}_1 = \{((x, y), (x', y')) \in \mathcal{V}_1 \times \mathcal{V}_1$ and $(x, y) \neq (x', y')\}$. The solution provides a coarse correspondence $\mathcal{C}_1 \subset X_{N_1} \times Y_{N_1}$.

At the $(k+1)$st level, correspondence is found between $X_{N_{k+1}}$ and $Y_{N_{k+1}}$ (the number of points is increased by a factor typically $2 \leq q = N_{k+1}/N_k \leq 4$), restricting the correspondence candidates for points within a certain radius around x to points within a certain radius around y, where $(x, y) \in \mathcal{C}_k$. This way, the $(k+1)$st level labeling problem is solved on the graph with vertices

$$\mathcal{V}_{k+1} = \{(x_i, y_i) \in X_{N_{k+1}} \times Y_{N_{k+1}} : \exists (x, y) \in \mathcal{C}_k \text{ s.t. } d_X(x, x_i) < \rho r_k,\ d_Y(y, y_i) < \rho r'_k\},$$

where $\rho > 1$, and $\mathcal{E}_{k+1} = \{((x, y), (x', y')) \in \mathcal{V}_{k+1} \times \mathcal{V}_{k+1}$ and $(x, y) \neq (x', y')\}$. For $\rho \approx 1$, the size of the ρr_k-neighborhood in $X_{N_{k+1}}$ of a point from X_{N_k} contains $\mathcal{O}(q)$ points. Thus, $|\mathcal{V}_{k+1}| = \mathcal{O}(q^2 N_k)$, and $|\mathcal{E}_{k+1}| = \mathcal{O}(q^4 N_k^2)$ points, a significant reduction compared to $\mathcal{O}(N_{k+1}^2)$ vertices and $\mathcal{O}(N_{k+1}^4)$ edges in a full graph, as shown Figure 1. As a result, the complexity of the optimization becomes $\mathcal{O}(N^4)$.

5 Probabilistic matching and shape prototypes

While invariance to geometric transformations such as bending can be accounted by the selection of local (descriptor) and global (metric) structures, many types of shape variability cannot be accounted for in this way. For example, variability within the

shape class (e.g. fat or thin man) results in different local and global structures that cannot be modeled explicitly. At the same time, such a variability can be modeled statistically. Instead of a descriptor and metric defining the shape X, we now have *distributions* $\mathbf{f}_X \sim \mathcal{F}_X$ and $d_X \sim \mathcal{D}_X$, e.g., *Gaussian mixture model* for distances,

$$\begin{aligned} \mathrm{p}_{xx'}(d) &= \sum_{k=1}^{K} \pi_{xx'k} \frac{1}{\sqrt{2\pi}\sigma_{xx'k}} \exp\left\{-\frac{(d-\mu_{xx'k})^2}{2\sigma_{xx'k}^2}\right\}, \\ & \sum_{k=1}^{K} \pi_{xx'k} = 1;\ \forall x \neq x' \in X, \end{aligned}$$

and descriptors,

$$\begin{aligned} \mathrm{p}_x(\mathbf{f}) &= \sum_{k=1}^{K} \pi_{xk} \frac{\exp\left\{-\frac{1}{2}(\mathbf{f}-\boldsymbol{\mu}_{xk})^{\mathrm{T}}\boldsymbol{\Sigma}_{xk}^{-1}(\mathbf{f}-\boldsymbol{\mu}_{xk})\right\}}{(2\pi)^{m/2}(\det \boldsymbol{\Sigma}_{xk})^{1/2}}, \\ & \sum_{i=1}^{K} \pi_{xk} = 1;\ \forall x \in X, \end{aligned}$$

where p denotes probability density. The distance distribution between points x and x' is parameterized by $\mathcal{D}_{xx'} = \{\mu_{xx'k}, \sigma_{xx'k}^2, \pi_{xx'k}\}_{k=1}^K$; the distribution of descriptors at each points x is parameterized by $\mathcal{F}_x = \{\boldsymbol{\mu}_{xk}, \boldsymbol{\Sigma}_{xk}, \pi_{xk}\}_{i=1}^K$, where $\boldsymbol{\mu}_{xk}$ are $m \times 1$ vectors and $\boldsymbol{\Sigma}_{xk}$ are $m \times m$ matrices. We call $\mathcal{X} = ((\mathcal{F}_x)_{x\in X}, (\mathcal{D}_{xx'})_{x\neq x'\in X})$ a *shape prototype*.

In this probabilistic setting, given a shape Y, we determine the correspondence between Y and the prototype $\mathcal{X}$ by solving a problem similar to (5), with the distortion terms replaced by negative log-likelihood functions,

$$\begin{aligned} \min_{u\in\mathcal{L}^{\mathcal{V}}} \quad & -\sum_{(x,y)\in\mathcal{V}} u_{x,y}(\log \mathrm{p}_x(\mathbf{f}_Y(y)) + \gamma) - \\ & \beta \sum_{((x,y),(x',y'))\in\mathcal{E}} u_{x,y}u_{x',y'} \log \mathrm{p}_{xx'}(d_Y(y,y')) \\ & \text{s.t.} \sum_y u_{x,y} \leq 1\ \ \forall x \in X; \quad \sum_x u_{x,y} \leq 1\ \ \forall y \in Y. \end{aligned} \tag{6}$$

6 Results

To assess the performance of the presented approach, we performed multiple experiments of shape correspondence and similarity computation under a variety of transformations. Shapes from the TOSCA [10] and Princeton [41] datasets were used in our experiments. Textured shapes acquired with a multicamera system were taken from the INRIA Grenoble dataset [48]. The shapes were represented as triangular meshes with 2000-10000 vertices. Textures were given as RGB values for each vertex.Geodesic distances were computed using fast marching [23]. Diffusion and commute time metrics were computed using the spectral formulae (3) and (4) taking the first 100 eigenvalues. The Laplace-Beltrami operator was approximated using cotangent weights [32] and its eigenvalues and eigenfunctions were computed by solving the generalized eigensystem as described in [25]. The heat kernel was approximated using formula (2). Hierarchical matching was implemented in MATLAB with discrete optimization module in C++.

Typical running times for pairwise shape matching in the following experiments were about $10 - 20$ sec.

6.1 Invariance and the choice of the metric/descriptor

In the first experiment, matching was performed between eight points with equal weight given to the local and global distortion terms in the optimization problem. Three combinations of first- and second-order structures were used: geodesic metric/HKS descriptor, diffusion metric/HKS descriptor, and commute time metric/SI-HKS descriptor. Figure 2 shows the result of correspondence computation between shapes with different transformations for different choice of metric/descriptor. All three methods are invariant to bendings (first row; note that correspondence is defined up to an intrinsic symmetry). The combination geodesic metric/HKS descriptor is sensitive to topology (a human with hands glued to legs, second row) and scale. The combination diffusion metric/HKS descriptor is insensitive to topology but sensitive to scale. Finally, commute time metric with SI-HKS descriptor are invariant to all of the above.

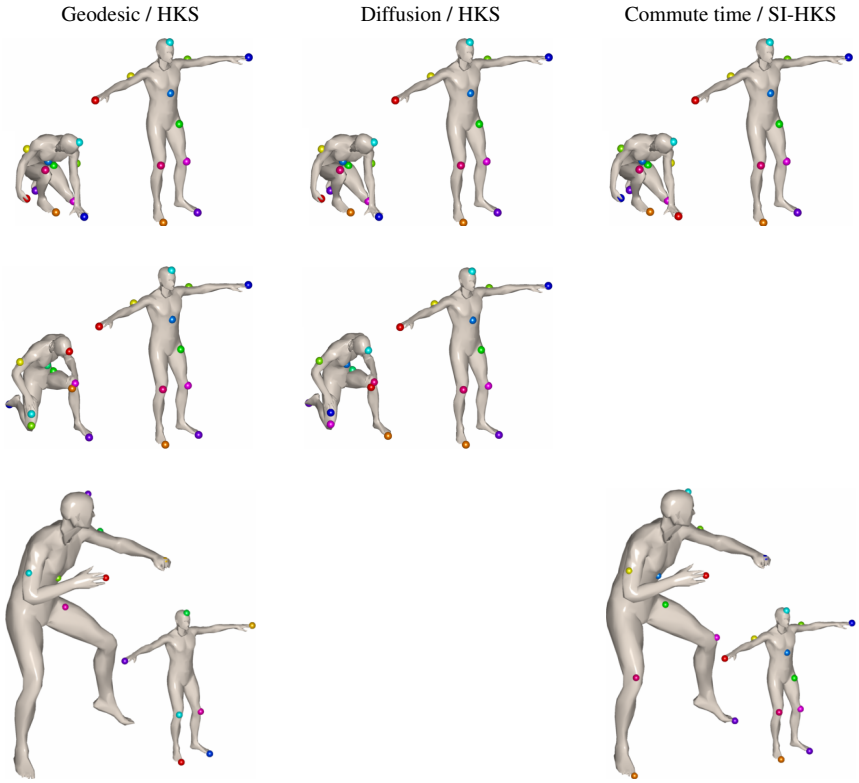


Figure 2: Invariance to different types of transformations and the choice of the metric/descriptor. Shown is matching between isometric deformations (first row), shapes with different topology (second row), and shapes with different scale (third row), using geodesic metric and HKS descriptors (left), diffusion metric and HKS descriptors (middle), and commute time metric and scale-invariant HKS descriptors (right).

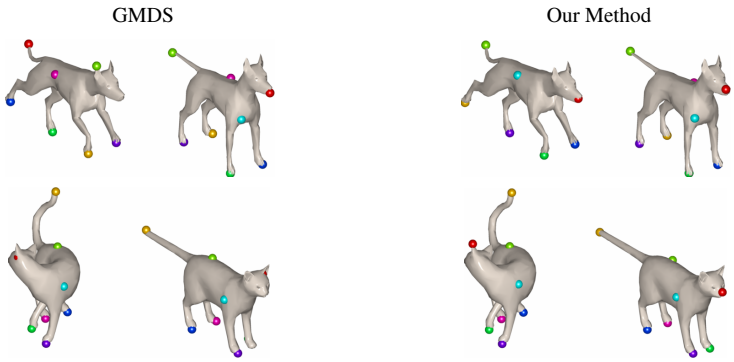


Figure 3: Comparison to GMDS.

6.2 Comparison to GMDS

In the second experiment, we compared our method to *generalized multidimensional scaling* (GMDS)[2], introduced in [8] and, as of today, a state-of-the-art method for deformable shape matching. GMDS computes correspondence between two shapes by trying to embed one shape into another with minimum distortion of geodesic distances (referred to as *stress*).

Given a fixed set of points $x_1, \ldots, x_N \in X$, GMDS attempts to find a set of corresponding points on Y in barycentric coordinates $y_i = (t_i, \mathbf{u}_i)$ (where $t_i \in T(Y)$ is a triangle index, and $u_{ij} \in [0,1], \sum_j u_{ij} = 1$ is a vector of barycentric weights) by minimizing the stress

$$\min_{\{t_i;\mathbf{u}_i\}_{i=1}^N} \sum_{i>j} (d_X(x_i, x_j) - d_Y((t_i;\mathbf{u}_i),(t_j;\mathbf{u}_j)))^2, \tag{7}$$

where $d_X(x_i, x_j)$ is a pre-computed geodesic distance between x_i and x_j on X, and d_Y is interpolated from pre-computed geodesic distances between points on Y.

While the stress function is highly non-convex, GMDS optimization is performed in a multi-resolution manner and in practice shows good convergence if initialized sufficiently close to the global minimizer [7]. A branch-and-bound technique with local histogram descriptors proposed in [37] to initialize the GMDS optimization was used in this experiment. For a fair comparison, we applied our method with geodesic distances in the second-order distortion term and no first-order term. The number of matched points was eight in both methods. Figure 3 shows the correspondence obtained using GMDS (left) and the proposed method (right). The first row shows a case when GMDS converges incorrectly. In the second row, the results of GMDS and the proposed method are similar.

Figure 4 shows similarity computed on a set of six shape classes, with four near-isometric deformations in each class (total 24 shapes) using GMDS and the proposed method. Similarity was computed as the distortion of correspondence between six corresponding points. Using similarity to rank matches in a retrieval experiment, in which the deformed shapes are used as queries, the mean average precision (mAP) achieved by our method is 100% (meaning that the first ranked matched is always from the correct shape class). For comparison, GMDS achieves only 40.1% mAP.

[2] Code taken from tosca.cs.technion.ac.il.

6.3 Matching of textured shapes

In the third experiment, to show the performance of our approach on the problem of textured shape matching from the INRIA Grenoble dataset [48]. These shapes contained various acquisition artifacts, such as topological noise and irregular triangulation. As local descriptors, we used the texture RGB values averaged in a small neighborhood of the point. Commute time distance was used as the second-order structure. Matching was performed between seven points, with equal weight given to the local and global distortion terms in the optimization problem. The matching results are shown in Figure 5.

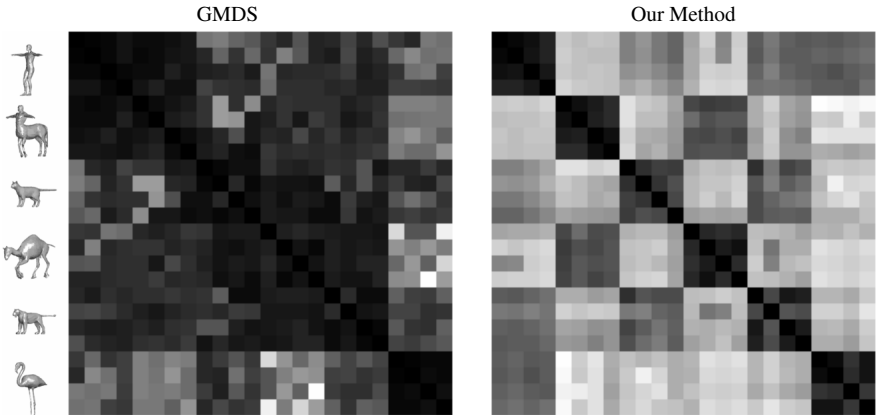


Figure 4: Dissimilarity between deformations of six classes of shapes computed using GMDS (left) the proposed method (right). Colormap encodes small dissimilarity values by darker colors. Ideal dissimilarity matrix has six 4×4 blocks of zeros on the diagonal and large values outside the diagonal.

Figure 5: Matching of textured surfaces using commute time metric and RGB texture.

6.4 Shape prototypes

In the fourth experiment, a shape prototype was created based on 64 examples of a human shape, in which the length of the hands and legs and the size of the head was varied. Distance and descriptor distributions were represented using Gaussian mixtures with 5 components. Figure 6 shows a comparison of deterministic and probabilistic matching. Using deterministic matching, the shape of a humanoid alien from the Princeton

database [41]. is matched to the human shape from TOSCA dataset incorrectly (second column from left), because of different proportions of the head, legs, and hands. On the other hand, matching to the human shape prototype using probabilistic matching produced correct symmetric correspondence (third column). Figure 6 (columns four and five) shows additional examples of shape prototype matching. These results show that the probabilistic matching framework allows to address shape variability that cannot be simply accommodated into the metric model by choosing the metric.

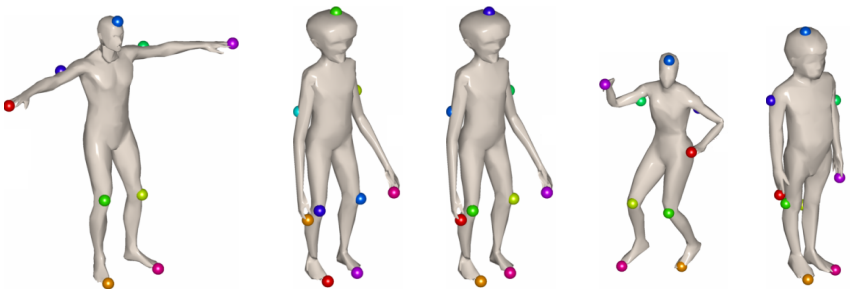


Figure 6: Matching of an alien shape to the human shape (first column from left) using deterministic (second column) and probabilistic (third column) approaches. Columns four and five: additional probabilistic matching examples.

7 Conclusions

We presented a generic framework for invariant matching between shapes, in which matching is performed by minimizing the distortion of local and global geometric structures under the correspondence. Using structures invariant to pre-defined classes of transformations (or, using their statistical distributions if such transformations cannot be modeled explicitly) allows obtaining invariant matching between shapes. Our approach generalizes many previous works in the field, in particular, methods based on metric distortion minimization [31, 8, 11] and global and local features [45, 16, 47], allowing incorporating many existing geometries and local descriptors [43, 14, 48, 1]. In particular, it extends the Gromov-Hausdorff framework [31, 8, 20]. Formulating the problem as graph labeling, we use powerful optimization method recently developed for this class of problems which are known to have favorable convergence properties. Our approach is especially appropriate for the challenging problems of finding similarity and correspondence between non-rigid shapes.

7.1 Limitations and extensions

The problem of symmetric correspondences, inherent to all approaches based on intrinsic structures, cannot be resolved without resorting to some side information. There are a few potential cures to this problem. First, providing some initial correspondence between the shapes could be used to restrict the vertex set, ruling out symmetric correspondences. Second, exploiting shape orientation could be used to find orientation-consistent matches. Finally, using higher-order distortions (in particular, third-order between triplets of points) can resolve the symmetry problem [49].

Though the presented approach shows better global convergence compared to GMDS, the advantage of GMDS is that it works with barycentric coordinates, thus finding correspondences at the sub-vertex resolution. This effect is especially significant in coarse meshes, where vertex-to-vertex correspondence can be too coarse. The proposed algorithm, can be used as an initialization to GMDS (which requires good initialization), e.g., replacing the branch-and-bound approach used in [37]. This way, GMDS would acts as a sub-vertex "refinement" stage.

References

[1] M. Ben-Chen, O. Weber, and C. Gotsman. Characterizing shape using conformal factors. In *Proc. Eurographics Workshop on Shape Retrieval*, 2008.

[2] A. Berg, T. Berg, and J. Malik. Shape matching and object recognition using low distortion correspondences. In *Proc. CVPR*, volume 1, 2005.

[3] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, 1999.

[4] S. Biasotti, B. Falcidieno, and M. Spagnuolo. Extended reeb graphs for surface understanding and description. *Lecture Notes in Computer Science*, pages 185–197, 2000.

[5] I. Borg and P. Groenen. *Modern multidimensional scaling - theory and applications*. Springer, 1997.

[6] A. M. Bronstein, M. M. Bronstein, L. J. Guibas, and M. Ovsjanikov. Shape Google: geometric words and expressions for invariant shape retrieval. *ACM Trans. Graphics*, 2010. to appear.

[7] A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Efficient computation of isometry-invariant distances between surfaces. *SIAM J. Sci. Comp.*, 28(5):1812–1836, 2006.

[8] A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Generalized multidimensional scaling: a framework for isometry-invariant partial surface matching. *PNAS*, 103(5):1168–1172, 2006.

[9] A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Calculus of non-rigid surfaces for geometry and texture manipulation. *Trans. Visualization and Computer Graphics*, 13(5):902–913, 2007.

[10] A. M. Bronstein, M. M. Bronstein, and R. Kimmel. *Numerical geometry of non-rigid shapes*. Springer, 2008.

[11] A. M. Bronstein, M. M. Bronstein, R. Kimmel, M. Mahmoudi, and G. Sapiro. A Gromov-Hausdorff framework with diffusion geometry for topologically-robust non-rigid shape matching. *IJCV*, 89(2–3):266–286, 2010.

[12] M. M. Bronstein and A. M. Bronstein. On a relation between shape recognition algorithms based on distributions of distances. Technical report, 2009.

[13] M. M. Bronstein and A. M. Bronstein. Shape recognition with spectral distances. *Trans. PAMI*, 2010. to appear.

[14] M. M. Bronstein and I. Kokkinos. Scale-invariant heat kernel signatures for shape recognition. In *Proc. CVPR*, 2010.

[15] R. R Coifman, S. Lafon, A. B. Lee, M. Maggioni, B. Nadler, F. Warner, and S. W. Zucker. Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. *PNAS*, 102(21):7426–7431, 2005.

[16] A. Dubrovina and R. Kimmel. Matching shapes by eigendecomposition of the Laplace-Beltrami operator. In *Proc. 3DPVT*, 2010.

[17] A. Elad and R. Kimmel. Bending invariant representations for surfaces. In *Proc. Computer Vision and Pattern Recognition (CVPR)*, pages 168–174, 2001.

[18] S. Gold and A. Rangarajan. A graduated assignment algorithm for graph matching. *Trans. PAMI*, 18:377–388, 1996.

[19] C. M. Grinstead and L. J. Snell. *Introduction to Probability*. AMS, 1998.

[20] M. Gromov. *Structures Métriques Pour les Variétés Riemanniennes*. Number 1 in Textes Mathématiques. 1981.

[21] D. Hochbaum and D. Shmoys. A best possible heuristic for the k-center problem. *Mathematics of Operations Research*, 10:2:180–184, 1985.

[22] J. Hu and J. Hua. Salient spectral geometric features for shape matching and retrieval. *Visual Computer*, 25(5):667–675, 2009.

[23] R. Kimmel and J. A. Sethian. Computing geodesic paths on manifolds. *PNAS*, 95(15):8431–8435, 1998.

[24] N. Komodakis, N. Paragios, and G. Tziritas. MRF optimization via dual decomposition: Message-passing revisited. In *Proc. ICCV*, 2007.

[25] B. Levy. Laplace-Beltrami eigenfunctions towards an algorithm that "understands" geometry. In *Proc. ICSMA*, 2006.

[26] Y. Lipman and T. Funkhouser. Möbius voting for surface correspondence. *TOG*, 28(3), 2009.

[27] D. Lowe. Distinctive image features from scale-invariant keypoint. *IJCV*, 2004.

[28] D. Mateus, R. P. Horaud, D. Knossow, F. Cuzzolin, and E. Boyer. Articulated shape matching using laplacian eigenfunctions and unsupervised point registration. *Proc. CVPR*, 2008.

[29] F. Mémoli. On the use of Gromov-Hausdorff distances for shape comparison. 2007.

[30] F. Mémoli. Spectral Gromov-Wasserstein distances for shape matching. In *Proc. NORDIA*, 2009.

[31] F. Mémoli and G. Sapiro. A theoretical and computational framework for isometry invariant recognition of point cloud data. *Foundations of Computational Mathematics*, 5:313–346, 2005.

[32] M. Meyer, M. Desbrun, P. Schroder, and A. H. Barr. Discrete differential-geometry operators for triangulated 2-manifolds. *Visualization and Mathematics III*, pages 35–57, 2003.

[33] N. J. Mitra, L. Guibas, J. Giesen, and M. Pauly. Probabilistic fingerprints for shapes. In *Proc. SGP*, 2006.

[34] M. Ovsjanikov, A. M. Bronstein, M. M. Bronstein, and L. J. Guibas. Shape Google: a computer vision approach to invariant shape retrieval. In *Proc. NORDIA*, 2009.

[35] M. Ovsjanikov, J. Sun, and L. J. Guibas. Global intrinsic symmetries of shapes. In *Computer Graphics Forum*, volume 27, pages 1341–1348, 2008.

[36] H. Qiu and E. R. Hancock. Clustering and embedding using commute times. *Trans. PAMI*, 29(11):1873–1890, 2007.

[37] D. Raviv, A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Symmetries of non-rigid shapes. In *Proc. NRTL*, 2007.

[38] D. Raviv, A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Full and partial symmetries of non-rigid shapes. *IJCV*, 89(1):18–39, 2010.

[39] D. Raviv, A. M. Bronstein, M. M. Bronstein, R. Kimmel, and G. Sapiro. Diffusion symmetries of non-rigid shapes. In *Proc. 3DPVT*, 2010.

[40] R. M. Rustamov. Laplace-Beltrami eigenfunctions for deformation invariant shape representation. In *Proc. SGP*, pages 225–233, 2007.

[41] P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser. The Princeton shape benchmark. In *Proc. SMI*, pages 167–178, 2004.

[42] J. Starck and A. Hilton. Correspondence labelling for widetimeframe free-form surface matching. In *Proc. ICCV*, 2007.

[43] J. Sun, M. Ovsjanikov, and L. J. Guibas. A concise and provably informative multi-scale signature based on heat diffusion. In *Computer Graphics Forum*, volume 28, pages 1383–1392, 2009.

[44] A. Tevs, M. Bokeloh, M. Wand, A. Schilling, and H.-P. Seidel. Isometric Registration of Ambiguous and Partial Data. *Computers & Graphics*, 32(2), 2008.

[45] N. Thorstensen and R. Keriven. Non-rigid Shape matching using Geometry and Photometry. *Proc. ACCV*, 2009.

[46] R. Toldo, U. Castellani, and A. Fusiello. Visual vocabulary signature for 3D object retrieval and partial matching. In *Proc. Eurographics Workshop on 3D Object Retrieval*, 2009.

[47] L. Torresani, V. Kolmogorov, and C. Rother. Feature correspondence via graph matching: Models and global optimization. In *Proc. ECCV*, pages 596–609, 2008.

[48] A. Zaharescu, E. Boyer, K. Varanasi, and R Horaud. Surface Feature Detection and Description with Applications to Mesh Matching. In *Proc. CVPR*, 2009.

[49] Y. Zeng, C. Wang, Y. Wang, X. Gu, D. Samaras, and N. Paragios. Dense non-rigid surface registration using high-order graph matching. In *Proc. CVPR*, 2010.

[50] E. Zhang, K. Mischaikow, and G. Turk. Feature-based surface parameterization and texture mapping. *TOG*, 24(1):1–27, 2005.

[51] H. Zhang, A. Sheffer, D. Cohen-Or, Q. Zhou, O. van Kaick, and A. Tagliasacchi. Deformation-driven shape correspondence. In *Computer Graphics Forum*, volume 27, pages 1431–1439, 2008.

Contents

Centre de recherche INRIA Saclay – Île-de-France
Parc Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 Orsay Cedex (France)

Centre de recherche INRIA Bordeaux – Sud Ouest : Domaine Universitaire - 351, cours de la Libération - 33405 Talence Cedex
Centre de recherche INRIA Grenoble – Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier
Centre de recherche INRIA Lille – Nord Europe : Parc Scientifique de la Haute Borne - 40, avenue Halley - 59650 Villeneuve d'Ascq
Centre de recherche INRIA Nancy – Grand Est : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex
Centre de recherche INRIA Paris – Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex
Centre de recherche INRIA Rennes – Bretagne Atlantique : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex
Centre de recherche INRIA Sophia Antipolis – Méditerranée : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399